

AD _____

Award Number DAMD17-94-J-4015

TITLE: Digital Image Database with Gold Standard and Performance Metrics for Mammographic Image Analysis Research

PRINCIPAL INVESTIGATOR: Kevin W. Bowyer, Ph.D.

CONTRACTING ORGANIZATION: University of South Florida
Tampa, Florida 33620

REPORT DATE: August 1999

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

DTIC QUALITY INSPECTED 4

20010216 073

| REPORT DOCUMENTATION PAGE | | | Form Approved OMB No. 0704-0188 | |
|--|--|-------------------------------|---|--|
| Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503. | | | | |
| 1. AGENCY USE ONLY (Leave blank) | | 2. REPORT DATE August 1999 | | 3. REPORT TYPE AND DATES COVERED Final (1 Jul 94 - 31 Jul 99) |
| 4. TITLE AND SUBTITLE Digital Image Database with Gold Standard and Performance Metrics for Mammographic Image Analysis Research | | | 5. FUNDING NUMBERS DAMD17-94-J-4015 | |
| 6. AUTHOR(S) Kevin W. Bowyer, Ph.D. | | | | |
| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of South Florida Tampa, Florida 33620 E*Mail: kwb@csee.usf.edu | | | 8. PERFORMING ORGANIZATION REPORT NUMBER | |
| 9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012 | | | 10. SPONSORING / MONITORING AGENCY REPORT NUMBER | |
| 11. SUPPLEMENTARY NOTES | | | | |
| 12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited | | | 12b. DISTRIBUTION CODE | |
| 13. ABSTRACT (Maximum 200 words) | | | | |

The Digital Database for Screening Mammography (DDSM) is an infrastructure resource for use by the mammogram image analysis research community. It consists of digitized images for screening mammogram cases, associated information and associated software tools. Approximately 2,500 cases of data are currently available for browsing on the world-wide-web. This report outlines the final state of the DDSM resource and reflects on problems encountered in the development of the resource.

| | | | | |
|---|--|---|---|--|
| 14. SUBJECT TERMS Breast Cancer | | | 15. NUMBER OF PAGES 18 | |
| | | | 16. PRICE CODE | |
| 17. SECURITY CLASSIFICATION OF REPORT Unclassified | 18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified | 19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified | 20. LIMITATION OF ABSTRACT Unlimited | |

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

____ Where copyrighted material is quoted, permission has been obtained to use such material.

____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

KWB Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

____ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

____ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

____ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

____ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

Kenn W. Bump 1 Sep 99
PI - Signature Date

Contents

| | | |
|---|--------------------------------------|----|
| 1 | Front Cover Page | 1 |
| 2 | Report Documentation Page (SF 298) | 2 |
| 3 | Foreword Page | 3 |
| 4 | Introduction | 5 |
| 5 | Current State of the DDSM Resource | 5 |
| 6 | Problems Encountered | 6 |
| 7 | Conclusions | 6 |
| A | List of personnel | 7 |
| B | Bibliography of related publications | 8 |
| C | Printed copy of DDSM main web page | 9 |
| D | Comments from users of the database | 15 |

4 Introduction

The Digital Database for Screening Mammography (DDSM) is an infrastructure resource for the mammogram image analysis research community. Its purpose is to make it possible for researchers to conduct a more rigorous experimental comparison of the performance of different image analysis techniques. This final report presents the current state of the DDSM resource, and summarizes some of the problems encountered in the work.

5 Current State of the DDSM Resource

The DDSM resource currently contains both image-related data and associated software tools. The image-related data is organized by case, where a “case” is the standard four images of a screening exam, plus information on patient age, a radiologist-specified BIRADS breast density rating, a radiologist-specified outline of the suspicious region(s) in an image, and a radiologist-specified subtlety rating for detection of the lesion(s) in the image.

The address for the DDSM web site is:

<http://marathon.csee.usf.edu/Mammography/Database.html>

A printed copy of this opening page of the web site appears as an appendix to this report. This opening page give a summary of the data available, and of the basic organization of the resource.

Approximately 2,500 cases of mammogram data can currently be browsed as image “thumbnails.” Approximately 200 additional cases of data will be released in the very near future, as soon as the quality control checks are completed. Data in DDSM comes from four clinical sites: Massachusetts General Hospital in Boston, Wake Forest University School of Medicine in North Carolina, Sacred Heart Hospital in Pensacola, Florida, and Washington University School of Medicine in Saint Louis. Three different types of digitizers were used at different sites at different times: DBA Systems, Lumisys, and Howtek. (The DBA scanner was used at MGH and was “retired” due to continuing performance difficulties.)

The search engine available on the web page allows the user to collect together thumbnails of all cases which satisfy a search query. The search query can be formulated in terms of the additional information associated with each case. This additional information includes BIRADS keywords for abnormality description, breast density rating on the BIRADS 1-to-4 scale, and other information. As example queries, a user could collect together and browse the thumbnails of all cancer cases which have clustered calcifications, or of all density rating 4 cases which have a spiculated lesion.

The software tools available on-line through the DDSM resource include a utility for viewing the images and image-related data, a routine for matching the results of a CAD detection program to the radiologist-specified ground-truth location of a lesion, and a pointer to a lossless JPEG image compression utility.

Thus in its final state, the DDSM resource will contain approximately 2,700 cases of data. The original goal was 3,000 cases of data. The shortfall is primarily due to higher-than-anticipated amounts of effort that had to be devoted to (1) re-digitization of images due to scanner-induced artifacts, and (2) manual inspection of images to guard against patient identifier information creeping through in the digitized images. Many cases of data were re-digitized

after quality control checks revealed the presence of unacceptable artifacts in the images. This problem is encountered at some level with all digitizers, but was most severe during the period that a DBA systems film digitizer was in use at MGH. In our experience, the Lumisys and Howtek digitizers had much lower levels of occurrence of unacceptable artifacts.

6 Problems Encountered

As documented in previous annual reports, we encountered various problems that were not originally anticipated. Of these, the numerous problems involved with the use of the DBA M2100 film digitizer were the most frustrating to the goals of the project. The primary negative effect of the DBA digitizer was that it introduced various types of unacceptable artifacts in the digitized images, necessitating a large number of re-digitizations. When it became clear to us that DBA was unable or unwilling to make the digitizer function as originally described, the DBA digitizer at MGH was replaced with a Howtek digitizer. Specific details of some of the problems encountered in attempting to use the DBA digitizer were described in the previous annual report, and are not repeated here.

Another initially unexpected problem was the "unavailability rate" of films (in particular, films from cancer cases) when requested from the archives. This was not a major problem, in that it was readily compensated for by collecting cases from a broader time frame than originally anticipated.

The problem of deleting all patient identifier information from the digitized images generated an enormous unanticipated workload in terms of manual inspection of digitized images. The problem traces to the fact that patient identifier information can appear in more than one place on a film, and in more than one form, and that the customs for this can vary between clinical sites. There is lettering placed on the film cassette at the time of image acquisition, typed gum labels placed on the film after the study is done, and, more rarely, handwritten information in the margin of a film. Some of these are not obvious when first visualizing the digitized image, but become readily apparent after image processing steps that might be a part of some CAD routines. As a result, we felt it necessary to introduce a manual inspection step in to check for, and digitally "black out," any patient identifier information in the digitized image.

7 Conclusions

At the time that this report is written, approximately 2,500 cases of data can be browsed on the web site. Approximately another 200 cases should be added to the DDSM resource in the very near future. Frequency of access to the database is continuing to grow, and many people are ordering tapes of the data set. We have received a number of positive comments about the database and the quality of the data. Some recent comments are reproduced in the an appendix. Based on such feedback, we feel that the database is already beginning to fulfill the goal of facilitating more rigorous research in mammogram image analysis.

A List of personnel

All of the following people were appointed to the grant account at USF for some period of time during the contract.

- Kevin Bowyer
- Kyong Chang
- Michael Heath
- Rong Li
- Maha Sallam
- Heather Taylor
- Shannon Thrasher
- Jian Wang
- Kevin Woods
- Naveen Yarlagada

B Bibliography of related publications

- Sallam, M.Y., and Bowyer, K.W. Registration and difference analysis of corresponding mammogram images, *Medical Image Analysis* 3 (2), 103-118, 1999.
- Woods, K.S. and Bowyer, K.W. Generating ROC curves for artificial neural networks, *IEEE Transactions on Medical Imaging*, 16 (3), 329-337, June 1997.
- Woods, K., Kegelmeyer, W.P., and Bowyer, K.W. Combination of multiple classifiers using local accuracy estimates, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19 (4), 405-410, April 1997.
- Heath, M.D. and Bowyer, K.W. Computer Aided Detection for Screening Mammography, in *The Image and Video Processing Handbook*, Al Bovik, editor, Academic Press, 1999.
- Woods, K.S., Bowyer, K.W., and Sallam, M.Y. Evaluating detection algorithms, chapter 3 in *Image Processing Techniques for Tumor Detection*, Robin Strickland, editor, Marcel Dekker publishers, 1999.
- Heath, M., Bowyer, K., Kopans, D., Kegelmeyer, P., Moore, R., Chang, K. and Munishkumaran, S. Current status of the Digital Database for Screening Mammography, pages 457-460 in *Digital Mammography*, Kluwer Academic Publishers, 1998. (proceedings of the Fourth International Workshop on Digital Mammography, Nijmegen, The Netherlands)
- Sallam, M.Y., Bowyer, K.W., Woods, K., Kopans, D.B., Moore, R.H., Kegelmeyer, W.P. The Digital Database for Screening Mammography (DDSM): lessons learned, *Radiological Society of North America (RSNA '97)*, Chicago, Illinois (December 1997).
- Bowyer, K.W., Kopans, D.B., Kegelmeyer, W.P., Moore, R.H., Sallam, M.Y., Chang, K., and Woods, K. The Digital Database for Screening Mammography Research, *DOD Breast Cancer Research Program Meeting: Era of Hope*, Washington DC, (October 1997), 69-70.
- Bowyer, K., Kopans, D., Kegelmeyer, W.P., Moore, R., Sallam, M., Chang, K., and Woods, K. The digital database for screening mammography, *Third International Workshop on Digital Mammography*, Chicago (June 1996), 431-434.
- Woods, K., and Bowyer, K. A general view of detection algorithms, *Third International Workshop on Digital Mammography*, Chicago (June 1996), 385-390.
- Sallam, M., and Bowyer, K. Detecting abnormal densities in mammograms by comparison to previous screenings, *Third International Workshop on Digital Mammography*, Chicago (June 1996), 417-420.



University of South Florida

Digital Mammography Home Page

DDSM: Digital Database for Screening Mammography

The Digital Database for Screening Mammography (DDSM) is a resource for use by the mammographic image analysis research community. Primary support for this project was a grant from the Breast Cancer Research Program of the U.S. Army Medical Research and Materiel Command. The DDSM project is a collaborative effort involving co-p.i.s at the Massachusetts General Hospital (D. Kopans, R. Moore), the University of South Florida (K. Bowyer), and Sandia National Laboratories (P. Kegelmeyer). Additional cases from Washington University School of Medicine were provided by Peter E. Shile, MD, Assistant Professor of Radiology and Internal Medicine. Additional collaborating institutions include Wake Forest University School of Medicine, Sacred Heart Hospital and ISMD, Incorporated. The primary purpose of the database is to facilitate sound research in the development of computer algorithms to aid in screening. Secondary purposes of the database may include the development of algorithms to aid in the diagnosis and the development of teaching or training aids. The database contains approximately 2,500 studies. Each study includes two images of each breast, along with some associated patient information (age at time of study, ACR breast density rating, subtly rating for abnormalities, ACR keyword description of abnormalities) and image information (scanner, spatial resolution, ...). Images containing suspicious areas have associated pixel-level "ground truth" information about the locations and types of suspicious regions. Also provided are software both for accessing the mammogram and truth images and for calculating performance figures for automated image analysis algorithms.

The Digital Database for Screening Mammography is organized into "cases" and "volumes." A "case" is a collection of images and information corresponding to one mammography exam of one patient. A "volume" is simply a collection of cases collected together for purposes of ease of distribution. The DDSM database is under construction. All volumes are available on 8mm tape, and at any given point in time, a number of volumes are also available on-line. The README file explaining "everything" about the database is available, and many answers to questions about the database are listed below.

- **What information is included in a case?**

A case consists of between 6 and 10 files. These are an "ics" file, an overview "16-bit PGM" file,

four image files that are compressed with lossless JPEG encoding and zero to four overlay files. Normal cases will not have any overlay files. Click here for more detailed information on the files contained in a case.

- **What is the difference between normal, cancer, benign and benign without callback volumes?**

Each volume is a collection of cases of the corresponding type. Normal cases are formed from a previous normal screening exam (pulled from a file) for a patient with a normal exam at least four years later. A normal screening exam is one in which no further "work-up" was required. Cancer cases are formed from screening exams in which at least one pathology proven cancer was found. Benign cases are formed from screening exams in which something suspicious was found, but was determined to not be malignant (by pathology, ultrasound or some other means). The term benign without callback is used to identify benign cases in which no additional films or biopsy was done to make the benign finding. These cases, however, contained something interesting enough for the radiologist to mark. A small number of cancer cases may contain, in addition to one or more regions that are path-proven malignant, one or more regions that are unproven. These are suspicious regions for which there is no path result. (Click here for more about ground truth.)

- **If I use data from DDSM in publications...**

Please credit the DDSM project as the source of the data, and reference "Current status of the Digital Database for Screening Mammography," M. Heath, K.W. Bowyer, D. Kopans et al, pages 457-460 in *Digital Mammography*, Kluwer Academic Publishers, 1998. Also, please send a copy of your publication to Professor Kevin Bowyer / Computer Science and Engineering / University of South Florida / Tampa, Florida 33620. We will eventually put a list of references on this web page.

- **What volumes are available?**

This database is still growing. The table below lists the volumes that are currently part of the database:

| VOLUME | CASES | SIZE | SCANNER | BITS | RESOLUTION | THUMBNAILS | NOTES | AVAI |
|-----------|-------|--------|---------|------|--------------|------------|-------|------|
| normal_01 | 111 | 5.8 GB | DBA | 16 | 42 microns | thumbnails | notes | |
| normal_02 | 117 | 6.6 GB | DBA | 16 | 42 microns | thumbnails | notes | |
| normal_03 | 38 | 4.1 GB | DBA | 16 | 42 microns | thumbnails | notes | |
| normal_04 | 57 | 5.1 GB | DBA | 16 | 42 microns | thumbnails | notes | |
| normal_05 | 47 | 4.3 GB | DBA | 16 | 42 microns | thumbnails | notes | |
| normal_06 | 60 | 5.5 GB | DBA | 16 | 42 microns | thumbnails | notes | |
| normal_07 | 78 | 6.2 GB | HOWTEK | 12 | 43.5 microns | thumbnails | notes | |

| | | | | | | | | |
|-----------|----|--------|---------|----|--------------|------------|-------|-----|
| normal_08 | 27 | 2.8 GB | HOWTEK | 12 | 43.5 microns | thumbnails | notes | |
| normal_09 | 59 | 4.9 GB | LUMYSIS | 12 | 50 microns | thumbnails | notes | |
| normal_10 | 23 | 2.1 GB | LUMYSIS | 12 | 50 microns | thumbnails | notes | |
| cancer_01 | 69 | 3.9 GB | LUMISYS | 12 | 50 microns | thumbnails | notes | |
| cancer_02 | 88 | 5.7 GB | LUMISYS | 12 | 50 microns | thumbnails | notes | |
| cancer_03 | 66 | 6.0 GB | DBA | 16 | 42 microns | thumbnails | notes | |
| cancer_04 | 31 | 2.8 GB | DBA | 16 | 42 microns | thumbnails | notes | |
| cancer_05 | 83 | 6.6 GB | LUMISYS | 12 | 50 microns | thumbnails | notes | |
| cancer_06 | 56 | 6.3 GB | HOWTEK | 12 | 43.5 microns | thumbnails | notes | |
| cancer_07 | 52 | 6.1 GB | HOWTEK | 12 | 43.5 microns | thumbnails | notes | |
| cancer_08 | 55 | 6.0 GB | HOWTEK | 12 | 43.5 microns | thumbnails | notes | |
| cancer_09 | 81 | 6.5 GB | LUMISYS | 12 | 50 microns | thumbnails | notes | |
| cancer_10 | 59 | 6.6 GB | HOWTEK | 12 | 43.5 microns | thumbnails | notes | ftf |
| cancer_11 | 59 | 5.9 GB | HOWTEK | 12 | 43.5 microns | thumbnails | notes | |
| cancer_12 | 83 | 6.8 GB | HOWTEK | 12 | 43.5 microns | thumbnails | notes | |
| cancer_13 | 18 | 1.5 GB | HOWTEK | 12 | 43.5 microns | thumbnails | notes | |
| benign_01 | 80 | 6.5 GB | LUMISYS | 12 | 50 microns | thumbnails | notes | |
| benign_02 | 69 | 6.9 GB | HOWTEK | 12 | 43.5 microns | thumbnails | notes | |
| benign_03 | 64 | 6.7 GB | HOWTEK | 12 | 43.5 microns | thumbnails | notes | |
| benign_04 | 81 | 6.5 GB | LUMISYS | 12 | 50 microns | thumbnails | notes | |
| benign_05 | 62 | 6.5 GB | HOWTEK | 12 | 43.5 microns | thumbnails | notes | |

| | | | | | | | | |
|--------------------|----|--------|---------|----|--------------|------------|-------|-----|
| benign_06 | 74 | 6.1 GB | LUMISYS | 12 | 50 microns | thumbnails | notes | |
| benign_07 | 61 | 6.1 GB | HOWTEK | 12 | 43.5 microns | thumbnails | notes | |
| benign_08 | 64 | 6.5 GB | HOWTEK | 12 | 43.5 microns | thumbnails | notes | ftp |
| benign_09 | 75 | 6.1 GB | HOWTEK | 12 | 43.5 microns | thumbnails | notes | |
| benign_10 | 21 | 2.1 GB | HOWTEK | 12 | 43.5 microns | thumbnails | notes | |
| benign_11 | 62 | 6.5 GB | HOWTEK | 12 | 43.5 microns | thumbnails | notes | |
| benign_12 | 64 | 6.4 GB | HOWTEK | 12 | 43.5 microns | thumbnails | notes | |
| benign_13 | 72 | 6.1 GB | LUMISYS | 12 | 50 microns | thumbnails | notes | |
| bwc_01 | 75 | 6.4 GB | LUMISYS | 12 | 50 microns | thumbnails | notes | |
| bwc = benign_witho | | | | | | | | |

● **Do you have a "troubleshooting" section on you web pages?**

Yes. We have compiled a list of frequently asked questions and have provided answers to them. Click here to go to that information.

● **How do I acquire a volume?**

Several volumes will be available by anonymous ftp at any given time (**figment.csee.usf.edu in pub/DDSM/cases**). You can download individual cases or entire volumes. Occasionally, we will change which volumes are available on line giving preference to the more recently released volumes. All volumes that are part of the database (whether they are on, or off line) can be ordered. Each is available on 8mm EXABYTE 160mXL data cartridges created using the UNIX tar command (and a model 8505XL 8mm drive). To order tapes, please specify the volume(s), and send a check of \$30.00 for the first tape plus \$20 for each additional tape. For international orders, add an additional \$20 for each three tapes ordered (1 to 3 tapes for \$20; 4 to 6 tapes for \$40 and so on). This is for customs and mailing. If we can find a cheaper way to do it, this may change in the future. Click here for an order form.

Make check payable to: **University of South Florida** (Please be careful that the check is not made out to "University of Florida", "Florida Southern University", "University of Southern Florida" or other variations; this can cause problems at the bank.)

Unfortunately, we are not set up to accept purchase orders or credit cards.

Checks must be made in U.S. dollars, drawn on a U.S. bank. Mail to:

Rachel Gadsden
University of South Florida
Department of Computer Science

4202 E. Fowler Ave.
ENB 118
Tampa, FL 33620-5399

- **What software is available for working with this data?**

We have software available for uncompressing image files, viewing cases, converting images to 16-BIT PGM format and utilities for comparing automated analysis results to ground truth. Source code from the Portable Video Research Group for the lossless JPEG compression program is available. Documentation on the use of the viewing software, DDSMView, is also available.

- **Can I preview the cases in a volume?**

Yes, we have made web pages that show "thumbnail" versions of the images. See the table for links to each volume of thumbnails. Each case has a separate web page. On each page, "thumbnail" images are displayed with all of the ground truth markings overlayed on them. The text information from the ics file and all of the overlay files is also provided. Please note that the colors for the overlayed ground truth markings are selected independently for each image. The color of each boundary can be used to index the associated textual information for that marking in the overlay table. Colors are not coordinated across MLO and CC views of the breast.

- **Can I search the cases in in the database?**

Yes, we have recently added a search capability to our database. [Click here to search the database.](#)

- **What is the "notes" link in the table of cases?**

The table of cases has a link to a page for each volume. Each page contains additional information about cases, such as presence of pacemaker, implants, skin markers, and other rare occurrences. The notes also contain information on any changes made to the cases after they were released. Although each case is checked thoroughly (and re-checked) before being released, errors may rarely exist in released volumes. When any errors are found, they will be corrected and listed on the notes page for that volume.

- **How do I map grey levels to optical density?**

In some situations, it may be useful to be able to map the grey levels in a mammogram image to optical density values. For example, you may want to run your image analysis software on data sets that were acquired on two different scanners. Since the grey levels in images acquired on different scanners will probably not correspond to the same optical density, you may want to "normalize" the images in some manner prior to processing them.

Here's how to map grey levels to optical density for images digitized at:

- DBA scanner
- HOWTEK scanner
- LUMISYS scanner

- **Are statistics available on patient population?**

The largest portion of the DDSM cases come from the Massachusetts General Hospital mammography program. Another substantial portion of the DDSM cases come from the Wake Forest University School of Medicine mammography program. All cases in DDSM are female patients, of course. The general statistical breakdown of patients by race at MGH and WFUSM is:

| MGH | WFUSM |
|-----|-------|
|-----|-------|

| | | |
|-----------------|-------|------|
| Asian | 2.06 | 0.2 |
| Black | 4.12 | 20.4 |
| Spanish Surname | 6.55 | 1.8 |
| American Indian | 0.00 | 0.1 |
| Other | 0.75 | 0.1 |
| Unknown | 30.34 | 0.3 |
| White | 56.18 | 77.0 |

- **How can I keep myself informed on updates/additions to this database?**

To place yourself on an electronic mailing list to receive updates about this project (including the eventual creation of mailing list discussion group), Click Here *Email: ddsm@bigpine.csee.usf.edu*
While we try to respond to technical questions directed to this email address, we DO NOT provide any clinical or patient advice. While we try to respond to technical questions in a timely manner, it may take a while for us to get back to you.

- **Do you have anonymous ftp access statistics available?**

Yes. We have a page displaying a graph showing the amount of data downloaded from DDSM (pub/DDSM/cases) by anonymous ftp each week. Click here to view the graph.

- **Are there other Mammography resources on this web site.**

Yes. They have been moved to our "Other Resources" page.

Note: The Digital Database for Screening Mammography (DDSM) is supported through a grant from the DOD Breast Cancer Research Program, US Army Research and Material Command DAMD17-94-J-4015. The server for the DDSM is a dual processor Sun Sparc 20 with 520 Megs of RAM donated by Sun Microsystems through their Academic Equipment Grant (AEG) program, Grant #: EDUD-US-950408.

Please mail comments, suggestions and specific mammography questions to:
ddsm@bigpine.csee.usf.edu

D Comments from users of the database

Date: Fri, 16 Jul 1999 21:44:30 -0400
From: James N <nguye02@med.mcgill.ca>
Subject: mammogram images
To: dds@bigpine.csee.usf.edu
Message-id: <002e01becff5\$eaf64da0\$d1b8a8c6@default>
MIME-version: 1.0
X-MIMEOLE: Produced By Microsoft MimeOLE V5.00.2615.200
X-Mailer: Microsoft Outlook Express 5.00.2615.200
X-Priority: 3
X-MSMail-priority: Normal
Content-Type: MULTIPART/ALTERNATIVE;
BOUNDARY="Boundary_(ID_oRRwOBoJacBCGjvH3vRm1g)"
Content-Length: 2410
Status: R

This is a multi-part message in MIME format.

--Boundary_(ID_oRRwOBoJacBCGjvH3vRm1g)
Content-type: text/plain; charset=iso-8859-1
Content-transfer-encoding: 7BIT

Dear Sir or Madam,

I am a medical student at McGill University in Montreal, Canada working on an internet project involving mammograms. We are planning to develop a teaching tutorial on a public internet website. We hope to make our site as interactive as possible in demonstrating the processes involved in analyzing mammograms. We are in the process of collecting our own images into cases. Your database of digitized mammogram images is impressive. Although we are not sure if we would incorporate any of the images within your database, we are inquiring as to whether you would grant us permission if the need arises. Our website is noncommercial and is intended for teaching purposes only. Thank you for your kind response and we look forward to hearing from you.

Sincerely,

James Nguyen

Date: Fri, 09 Jul 1999 21:32:30 -0700
To: Dr Kevin Bowyer <kwb@bigpine.csee.usf.edu>
From: Jack Sklansky <sklansky@uci.edu>
Subject: your request for PAMI.
In-Reply-To: <199907091359.JAA17230@keylime.csee.usf.edu>
Mime-Version: 1.0
Content-Type: text/plain; charset="us-ascii"
Content-Length: 512
Status: R

Dear Kevin --

I am honored by your request. Unfortunately, I just finished reviewing a manuscript for another journal, and a second manuscript is awaiting my review. So I must decline your request.

By the way -- last week we began organizing mammographic data that I received from your institution. The data will be used in a test of our CAD system. We are very pleased with the quality of this data. Congratulations on producing this database.

-- Jack

=====

Message-ID: <000c01bec1a6\$5abfe550\$0500a8c0@paul>
From: "paul" <paul@gmai.com>
To: "Dr Kevin Bowyer" <kwb@bigpine.csee.usf.edu>
Cc: "Ernest Keenan" <ern@gmai.com>
References: <199906281709.NAA04591@keylime.csee.usf.edu>
Subject: Re: discussion
Date: Mon, 28 Jun 1999 13:39:42 -0700
MIME-Version: 1.0
Content-Transfer-Encoding: 7bit
X-Priority: 3
X-MSMail-Priority: Normal
X-Mailer: Microsoft Outlook Express 5.00.2314.1300
X-MimeOLE: Produced By Microsoft MimeOLE V5.00.2314.1300
Content-Type: text/plain;
charset="iso-8859-1"
Content-Length: 2413
Status: RO

I am just starting a project funded by the National Cancer Institute. We will develop a data mining/ data warehouse / data store approach to medical image understanding... using situational logics and neural networks.

I will have my initial research notes placed in a web site by the end of the week. These will be updated and perhaps contributed to by others.

My work on situational logics and machine intelligence is at

www.bcnngroup.org

I do not have a specific question about your database. It is a fine resource, and I am grateful that it is available for research purposes. We already have a few of your tapes and will eventually want to acquire specific subcollections in several planned studies of how a machine intelligence architecture

<http://www.bcnngroup.org/area3/pprueitt/book.htm>

works with respect to images.

Specifically we are looking at Hopfield networks, and wavelets, to find and encode complex memories of feature constraints, related to multiscale resolution. First I need to develop an enumeration of pixel patterns found in localized images of abnormalities (seen at the same spacial resolution.) We are using a small data base found at

<http://peipa.essex.ac.uk/ipa/pix/mias/README> , for the initial study.

Later I plan to develop an extensive library of same resolution, same window size, images of mass abnormalities. These will be placed into a data warehouse for automated review and analysis (I think).

What I do hope to do is to provide some collaboration tools, that I have already developed, to the medical image understanding research community... and publish a few papers.

My area is neuropsychology, mathematics and logic.... with only limited awareness of the issues related specifically to mammography interpretation.

I am always very aware of the time constraints we are under....., and will try to make my communications with you and your group focused and meaningful.